



#4

SEQUENCE LISTING

<110> CALIFORNIA INSTITUTE OF TECHNOLOGY
Debe, Derek A.

<120> METHOD FOR DETERMINING THREE-DIMENSIONAL PROTEIN STRUCTURE FROM PRIMARY PROTEIN SEQUENCE

<130> 265/297

<140> US 09/905,176

<141> 2001-07-12

<150> US 60/218,016

<151> 2000-07-12

<160> 26

<170> PatentIn version 3.1

<210> 1

<211> 53

<212> PRT

<213> Artificial Sequence

<220>

<223> Genus/species, Unknown

<400> 1

Leu Val Ala Phe Ala Asp Phe Gly Ser Val Thr Phe Thr Asn Ala Glu
1 5 10 15

Ala Thr Ser Gly Gly Ser Thr Val Gly Pro Ser Asp Ala Thr Val Met
20 25 30

Asp Ile Glu Gln Asp Gly Ser Val Leu Thr Glu Thr Ser Val Ser Gly
35 40 45

Asp Ser Val Thr Val
50

<210> 2
<211> 53
<212> PRT
<213> Artificial Sequence

<220>
<223> Genus/species, Unknown

<400> 2

Leu Val Ala Phe Ala Asp Phe Gly Ser Val Thr Phe Thr Asn Ala Glu
1 5 10 15

Ala Thr Ser Gly Gly Ser Thr Val Gly Pro Ser Asp Ala Thr Val Met
20 25 30

Asp Ile Glu Gln Asp Gly Ser Val Leu Thr Glu Thr Ser Val Ser Gly
35 40 45

Asp Ser Val Thr Val
50

<210> 3
<211> 53
<212> PRT
<213> Artificial Sequence

<220>
<223> Genus/species, Unknown

<400> 3

Leu Val Pro Phe Ala Asn Phe Gly Thr Val Thr Phe Thr Gly Ala Glu
1 5 10 15

Ala Thr Thr Ser Ser Gly Thr Val Thr Ala Ala Asp Ala Thr Leu Ile
20 25 30

Asp Ile Glu Gln Asn Gly Glu Val Leu Thr Ser Val Thr Val Ser Gly
35 40 45

Ser Thr Val Thr Val
50

<210> 4
<211> 52
<212> PRT
<213> Artificial Sequence

<220>

<223> Genus/species, Unknown

<400> 4

Leu Val Gln Phe Ala Asn Phe Gly Thr Val Thr Phe Thr Gly Ala Ser
1 5 10 15

Ala Thr Gln Asn Gly Glu Ser Val Gly Val Thr Gly Ala Gln Ile Ile
20 25 30

Asp Leu Gln Gln Asn Ser Val Leu Thr Ser Val Ser Thr Ser Ser Asn
35 40 45

Ser Val Thr Val
50

<210> 5

<211> 47

<212> PRT

<213> Artificial Sequence

<220>

<223> Genus/species, Unknown

<400> 5

Leu Val Asn Phe Ala Asp Phe Asp Thr Val Thr Phe Lys Asp Cys Ser
1 5 10 15

Pro Ser Val Ser Gly Ser Thr Ile Val Asp Ile Arg Gln Ser Leu Glu
20 25 30

Val Leu Thr Glu Cys Ser Thr Thr Gly Thr Thr Thr Val Thr Cys
35 40 45

<210> 6

<211> 54

<212> PRT

<213> Artificial Sequence

<220>

<223> Genus/species, Unknown

<400> 6

Phe Val Pro Phe Ala Ser Phe Ser Pro Ala Val Glu Phe Thr Asp Cys
1 5 10 15

Ser Val Thr Ser Asp Gly Glu Ser Val Ser Leu Asp Asp Ala Gln Ile
20 25 30

Thr Gln Val Ile Ile Asn Asn Gln Asp Val Thr Asp Cys Ser Val Ser
35 40 45

Gly Thr Thr Val Ser Cys
50

<210> 7
<211> 54
<212> PRT
<213> Artificial Sequence

<220>
<223> Genus/species, Unknown

<400> 7

Phe Val Pro Phe Ala Ser Phe Ser Pro Ala Val Glu Phe Thr Asp Cys
1 5 10 15

Ser Val Thr Ser Asp Gly Glu Ser Val Ser Leu Asp Asp Ala Gln Ile
20 25 30

Thr Gln Val Ile Ile Asn Asn Gln Asp Val Thr Asp Cys Ser Val Ser
35 40 45

Gly Thr Thr Val Ser Cys
50

<210> 8
<211> 54
<212> PRT
<213> Pseudomonas aeruginosa

<400> 8

Phe Val Pro Phe Ala Ser Phe Ser Pro Ala Val Glu Phe Thr Asp Cys
1 5 10 15

Ser Val Thr Ser Asp Gly Glu Ser Val Ser Leu Asp Asp Ala Gln Ile
20 25 30

Thr Gln Val Ile Ile Asn Asn Gln Asp Val Thr Asp Cys Ser Val Ser
35 40 45

Gly Thr Thr Val Ser Cys
50

<210> 9
<211> 326
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 9

Tyr Pro Tyr Thr Arg Leu Arg Arg Asn Arg Arg Asp Asp Phe Ser Arg
1 5 10 15

Arg Leu Val Arg Glu Asn Val Leu Thr Val Asp Asp Leu Ile Leu Pro
20 25 30

Val Phe Val Leu Asp Gly Val Asn Gln Arg Glu Ser Ile Pro Ser Met
35 40 45

Pro Gly Val Glu Arg Leu Ser Ile Asp Gln Leu Leu Ile Glu Ala Glu
50 55 60

Glu Trp Val Ala Leu Gly Ile Pro Ala Leu Ala Leu Phe Pro Val Thr
65 70 75 80

Pro Val Glu Lys Lys Ser Leu Asp Ala Ala Glu Ala Tyr Asn Pro Glu
85 90 95

Gly Ile Ala Gln Arg Ala Thr Arg Ala Leu Arg Glu Arg Phe Pro Glu
100 105 110

Leu Gly Ile Ile Thr Asp Val Ala Leu Asp Pro Phe Thr Thr His Gly
115 120 125

Gln Asp Gly Ile Leu Asp Asp Asp Gly Tyr Val Leu Asn Asp Val Ser
130 135 140

Ile Asp Val Leu Val Arg Gln Ala Leu Ser His Ala Glu Ala Gly Ala
145 150 155 160

Gln Val Val Ala Pro Ser Asp Met Met Asp Gly Arg Ile Gly Ala Ile
165 170 175

Arg Glu Ala Leu Glu Ser Ala Gly His Thr Asn Val Arg Ile Met Ala
180 185 190

Tyr Ser Ala Lys Tyr Ala Ser Ala Tyr Tyr Gly Pro Phe Arg Asp Ala
195 200 205

Val Gly Ser Ala Ser Asn Leu Gly Lys Gly Asn Lys Ala Thr Tyr Gln
210 215 220

Met Asp Pro Ala Asn Ser Asp Glu Ala Leu His Glu Val Ala Ala Asp
225 230 235 240

Leu Ala Glu Gly Ala Asp Met Val Met Val Lys Pro Gly Met Pro Tyr
245 250 255

Leu Asp Ile Val Arg Arg Val Lys Asp Glu Phe Arg Ala Pro Thr Phe
260 265 270

Val Tyr Gln Val Ser Gly Glu Tyr Ala Met His Met Gly Ala Ile Gln
275 280 285

Asn Gly Trp Leu Ala Glu Ser Val Ile Leu Glu Ser Leu Thr Ala Phe
290 295 300

Lys Arg Ala Gly Ala Asp Gly Ile Leu Thr Tyr Phe Ala Lys Gln Ala
305 310 315 320

Ala Glu Gln Leu Arg Arg
325

<210> 10
<211> 328
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 10

Glu Ile Ser Ser Val Leu Ala Gly Gly Tyr Asn His Pro Leu Leu Arg
1 5 10 15

Gln Trp Gln Ser Glu Arg Gln Leu Thr Lys Asn Met Leu Ile Phe Pro
20 25 30

Leu Phe Ile Ser Asp Asn Pro Asp Asp Phe Thr Glu Ile Asp Ser Leu
35 40 45

Pro Asn Ile Asn Arg Ile Gly Val Asn Arg Leu Lys Asp Tyr Leu Lys
50 55 60

Pro Leu Val Ala Lys Gly Leu Arg Ser Val Ile Leu Phe Gly Val Pro
65 70 75 80

Leu Ile Pro Gly Thr Lys Asp Pro Val Gly Thr Ala Ala Asp Asp Pro
85 90 95

Ala Gly Pro Val Ile Gln Gly Ile Lys Phe Ile Arg Glu Tyr Phe Pro
100 105 110

Glu Leu Tyr Ile Ile Cys Asp Val Cys Leu Cys Glu Tyr Thr Ser His
115 120 125

Gly His Cys Gly Val Leu Tyr Asp Asp Gly Thr Ile Asn Arg Glu Arg
130 135 140

Ser Val Ser Arg Leu Ala Ala Val Ala Val Asn Tyr Ala Lys Ala Gly
145 150 155 160

Ala His Cys Val Ala Pro Ser Asp Met Ile Asp Gly Arg Ile Arg Asp
165 170 175

Ile Lys Arg Gly Leu Ile Asn Ala Asn Leu Ala His Lys Thr Phe Val
180 185 190

Leu Ser Tyr Ala Ala Lys Phe Ser Gly Asn Leu Tyr Gly Pro Phe Arg
195 200 205

Asp Ala Ala Cys Ser Ala Pro Ser Asn Gly Asp Arg Lys Cys Tyr Gln
210 215 220

Leu Pro Pro Ala Gly Arg Gly Leu Ala Arg Arg Ala Leu Glu Arg Asp
225 230 235 240

Met Ser Glu Gly Ala Asp Gly Ile Ile Val Lys Pro Ser Thr Phe Tyr
245 250 255

Leu Asp Ile Met Arg Asp Ala Ser Glu Ile Cys Lys Asp Leu Pro Ile
260 265 270

Cys Ala Tyr His Val Ser Gly Glu Tyr Ala Met Leu His Ala Ala Ala

275

280

285

Glu Lys Gly Val Val Asp Leu Lys Thr Ile Ala Phe Glu Ser His Gln
290 295 300

Gly Phe Leu Arg Ala Gly Ala Arg Leu Ile Ile Thr Tyr Leu Ala Pro
305 310 315 320

Glu Phe Leu Asp Trp Leu Asp Glu
325

<210> 11
<211> 215
<212> PRT
<213> Homo sapiens

<400> 11

Ala Gly His Asp Asn Thr Lys Pro Asp Thr Ser Ser Ser Leu Leu Thr
1 5 10 15

Ser Leu Asn Gln Leu Gly Glu Arg Gln Leu Leu Ser Val Val Lys Trp
20 25 30

Ser Lys Ser Leu Pro Gly Phe Arg Asn Leu His Ile Asp Asp Gln Ile
35 40 45

Thr Leu Ile Gln Tyr Ser Trp Met Ser Leu Met Val Phe Gly Leu Gly
50 55 60

Trp Arg Ser Tyr Lys His Val Ser Gly Gln Met Leu Tyr Phe Ala Pro
65 70 75 80

Asp Leu Ile Leu Asn Glu Gln Arg Met Lys Glu Ser Ser Phe Tyr Ser
85 90 95

Leu Cys Leu Thr Met Trp Gln Ile Pro Gln Glu Phe Val Lys Leu Gln
100 105 110

Val Ser Gln Glu Glu Phe Leu Cys Met Lys Val Leu Leu Leu Asn
115 120 125

Thr Ile Pro Leu Glu Gly Leu Arg Ser Gln Thr Gln Phe Glu Glu Met
130 135 140

Arg Ser Ser Tyr Ile Arg Glu Leu Ile Lys Ala Ile Gly Leu Arg Gln
145 150 155 160

Lys Gly Val Val Ser Ser Ser Gln Arg Phe Tyr Gln Leu Thr Lys Leu
165 170 175

Leu Asp Asn Leu His Asp Leu Val Lys Gln Leu His Leu Tyr Cys Leu
180 185 190

Asn Thr Phe Ile Gln Ser Arg Ala Leu Ser Val Glu Phe Pro Glu Met
195 200 205

Met Ser Glu Val Ile Ala Ala
210 215

<210> 12
<211> 207
<212> PRT
<213> Homo sapiens

<400> 12

Glu Ala Asn Met Gly Leu Asn Pro Ser Ser Pro Asn Asp Pro Val Thr
1 5 10 15

Asn Ile Cys Gln Ala Ala Asp Lys Gln Leu Phe Thr Leu Val Glu Trp
20 25 30

Ala Lys Arg Ile Pro His Phe Ser Glu Leu Pro Leu Asp Asp Gln Val
35 40 45

Ile Leu Leu Arg Ala Gly Trp Asn Glu Leu Leu Ile Ala Ser Ala Ser
50 55 60

His Arg Ser Ile Ala Val Lys Asp Gly Ile Leu Leu Ala Thr Gly Leu
65 70 75 80

His Val His Arg Asn Ser Ala His Ser Ala Gly Val Gly Ala Ile Phe
85 90 95

Asp Arg Val Leu Thr Glu Leu Val Ser Lys Met Arg Asp Met Gln Met
100 105 110

Asp Lys Thr Glu Leu Gly Cys Leu Arg Ala Ile Val Leu Phe Asn Pro

115

120

125

Asp Ser Lys Gly Leu Ser Asn Pro Ala Glu Val Glu Ala Leu Arg Glu
130 135 140

Lys Val Tyr Ala Ser Leu Glu Ala Tyr Cys Lys His Lys Tyr Pro Glu
145 150 155 160

Gln Pro Gly Arg Phe Ala Lys Leu Leu Arg Leu Pro Ala Leu Arg
165 170 175

Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Lys Leu Ile
180 185 190

Gly Asp Thr Pro Ile Asp Thr Phe Leu Met Glu Met Leu Glu Ala
195 200 205

<210> 13

<211> 240

<212> PRT

<213> Homo sapiens

<400> 13

Gln Leu Ile Pro Pro Leu Ile Asn Leu Leu Met Ser Ile Glu Pro Asp
1 5 10 15

Val Ile Tyr Ala Gly His Asp Asn Thr Lys Pro Asp Thr Ser Ser Ser
20 25 30

Leu Leu Thr Ser Leu Asn Gln Leu Gly Glu Arg Gln Leu Leu Ser Val
35 40 45

Val Lys Trp Ser Lys Ser Leu Pro Gly Phe Arg Asn Leu His Ile Asp
50 55 60

Asp Gln Ile Thr Leu Ile Gln Tyr Ser Trp Met Ser Leu Met Val Phe
65 70 75 80

Gly Leu Gly Trp Arg Ser Tyr Lys His Val Ser Gly Gln Met Leu Tyr
85 90 95

Phe Ala Pro Asp Leu Ile Leu Asn Glu Gln Arg Met Lys Glu Ser Ser
100 105 110

Phe Tyr Ser Leu Cys Leu Thr Met Trp Gln Ile Pro Gln Glu Phe Val
115 120 125

Lys Leu Gln Val Ser Gln Glu Glu Phe Leu Cys Met Lys Val Leu Leu
130 135 140

Leu Leu Asn Thr Ile Pro Leu Glu Gly Leu Arg Ser Gln Thr Gln Phe
145 150 155 160

Glu Glu Met Arg Ser Ser Tyr Ile Arg Glu Leu Ile Lys Ala Ile Gly
165 170 175

Leu Arg Gln Lys Gly Val Val Ser Ser Ser Gln Arg Phe Tyr Gln Leu
180 185 190

Thr Lys Leu Leu Asp Asn Leu His Asp Leu Val Lys Gln Leu His Leu
195 200 205

Tyr Cys Leu Asn Thr Phe Ile Gln Ser Arg Ala Leu Ser Val Glu Phe
210 215 220

Pro Glu Met Met Ser Glu Val Ile Ala Ala Gln Leu Pro Lys Ile Leu
225 230 235 240

<210> 14
<211> 241
<212> PRT
<213> Homo sapiens

<400> 14

Leu Thr Ala Asp Gln Met Val Ser Ala Leu Leu Asp Ala Glu Pro Pro
1 5 10 15

Ile Leu Tyr Ser Glu Tyr Asp Pro Thr Arg Pro Phe Ser Glu Ala Ser
20 25 30

Met Met Gly Leu Leu Thr Asn Leu Ala Asp Arg Glu Leu Val His Met
35 40 45

Ile Asn Trp Ala Lys Arg Val Pro Gly Phe Val Asp Leu Thr Leu His
50 55 60

Asp Gln Val His Leu Leu Glu Cys Ala Trp Leu Glu Ile Leu Met Ile

65

70

75

80

Gly Leu Val Trp Arg Ser Met Glu His Pro Gly Lys Leu Leu Phe Ala
85 90 95

Pro Asn Leu Leu Leu Asp Arg Asn Gln Gly Lys Cys Val Glu Gly Met
100 105 110

Val Glu Ile Phe Asp Met Leu Leu Ala Thr Ser Ser Arg Phe Arg Met
115 120 125

Met Asn Leu Gln Gly Glu Glu Phe Val Cys Leu Lys Ser Ile Ile Leu
130 135 140

Leu Asn Ser Gly Val Tyr Thr Phe Leu Ser Ser Thr Leu Lys Ser Leu
145 150 155 160

Glu Glu Lys Asp His Ile His Arg Val Leu Asp Lys Ile Thr Asp Thr
165 170 175

Leu Ile His Leu Met Ala Lys Ala Gly Leu Thr Leu Gln Gln Gln His
180 185 190

Glu Arg Leu Ala Gln Leu Leu Ile Leu Ser His Ile Arg His Met
195 200 205

Ser Asn Lys Gly Met Glu His Leu Tyr Ser Met Lys Cys Lys Asn Val
210 215 220

Val Pro Leu Tyr Asp Leu Leu Leu Glu Met Leu Asp Ala His Arg Leu
225 230 235 240

His

<210> 15

<211> 222

<212> PRT

<213> Halobacterium salinarum

<400> 15

Thr Gly Arg Pro Glu Trp Ile Trp Leu Ala Leu Gly Thr Ala Leu Met
1 5 10 15

Gly Leu Gly Thr Leu Tyr Phe Leu Val Lys Gly Met Gly Val Ser Asp
20 25 30

Pro Asp Ala Lys Lys Phe Tyr Ala Ile Thr Thr Leu Val Pro Ala Ile
35 40 45

Ala Phe Thr Met Tyr Leu Ser Met Leu Leu Gly Tyr Gly Leu Thr Met
50 55 60

Val Pro Phe Gly Gly Glu Gln Asn Pro Ile Tyr Trp Ala Arg Tyr Ala
65 70 75 80

Asp Trp Leu Phe Thr Thr Pro Leu Leu Leu Asp Leu Ala Leu Leu
85 90 95

Val Asp Ala Asp Gln Gly Thr Ile Leu Ala Leu Val Gly Ala Asp Gly
100 105 110

Ile Met Ile Gly Thr Gly Leu Val Gly Ala Leu Thr Lys Val Tyr Ser
115 120 125

Tyr Arg Phe Val Trp Trp Ala Ile Ser Thr Ala Ala Met Leu Tyr Ile
130 135 140

Leu Tyr Val Leu Phe Phe Gly Phe Ser Met Arg Pro Glu Val Ala Ser
145 150 155 160

Thr Phe Lys Val Leu Arg Asn Val Thr Val Val Leu Trp Ser Ala Tyr
165 170 175

Pro Val Val Trp Leu Ile Gly Ser Glu Gly Ala Gly Ile Val Pro Leu
180 185 190

Asn Ile Glu Thr Leu Leu Phe Met Val Leu Asp Val Ser Ala Lys Val
195 200 205

Gly Phe Gly Leu Ile Leu Leu Arg Ser Arg Ala Ile Phe Gly
210 215 220

<210> 16
<211> 233
<212> PRT
<213> Halobacterium salinarum

<400> 16

Glu Asn Ala Leu Leu Ser Ser Ser Leu Trp Val Asn Val Ala Leu Ala
1 5 10 15

Gly Ile Ala Ile Leu Val Phe Val Tyr Met Gly Arg Thr Ile Arg Pro
20 25 30

Gly Arg Pro Arg Leu Ile Trp Gly Ala Thr Leu Met Ile Pro Leu Val
35 40 45

Ser Ile Ser Ser Tyr Leu Gly Leu Leu Ser Gly Leu Thr Val Gly Met
50 55 60

Ile Glu Met Pro Ala Gly His Ala Leu Ala Gly Glu Met Val Arg Ser
65 70 75 80

Gln Trp Gly Arg Tyr Leu Thr Trp Ala Leu Ser Thr Pro Met Ile Leu
85 90 95

Leu Ala Leu Gly Leu Leu Ala Asp Val Asp Leu Gly Ser Leu Phe Thr
100 105 110

Val Ile Ala Ala Asp Ile Gly Met Cys Val Thr Gly Leu Ala Ala Ala
115 120 125

Met Thr Thr Ser Ala Leu Leu Phe Arg Trp Ala Phe Tyr Ala Ile Ser
130 135 140

Cys Ala Phe Phe Val Val Val Leu Ser Ala Leu Val Thr Asp Trp Ala
145 150 155 160

Ala Ser Ala Ser Ser Ala Gly Thr Ala Glu Ile Phe Asp Thr Leu Arg
165 170 175

Val Leu Thr Val Val Leu Trp Leu Gly Tyr Pro Ile Val Trp Ala Val
180 185 190

Gly Val Glu Gly Leu Ala Leu Val Gln Ser Val Gly Ala Thr Ser Trp
195 200 205

Ala Tyr Ser Val Leu Asp Val Phe Ala Lys Tyr Val Phe Ala Phe Ile
210 215 220

Leu Leu Arg Trp Val Ala Asn Asn Glu
225 230

<210> 17
<211> 267
<212> PRT
<213> Unknown

<220>
<223> Bovine Rhodopsin, 1+88A, Palczewski, K. et al., A G-protein Coupled Receptor, 289 Science 739 (2000)

<400> 17

Phe Ser Met Leu Ala Ala Tyr Met Phe Leu Leu Ile Met Leu Gly Phe
1 5 10 15

Pro Ile Asn Phe Leu Thr Leu Tyr Val Thr Val Gln His Lys Lys Leu
20 25 30

Arg Thr Pro Leu Asn Tyr Ile Leu Leu Asn Leu Ala Val Ala Asp Leu
35 40 45

Phe Met Phe Gly Gly Phe Thr Thr Thr Leu Tyr Thr Ser Leu His Gly
50 55 60

Tyr Phe Val Phe Gly Pro Thr Gly Cys Asn Leu Glu Gly Phe Phe Ala
65 70 75 80

Thr Leu Gly Gly Glu Ile Ala Leu Trp Ser Leu Val Val Leu Ala Ile
85 90 95

Glu Arg Tyr Val Val Val Cys Lys Pro Met Ser Asn Phe Arg Phe Gly
100 105 110

Glu Asn His Ala Ile Met Gly Val Ala Phe Thr Trp Val Met Ala Leu
115 120 125

Ala Cys Ala Ala Pro Pro Leu Val Gly Trp Ser Arg Tyr Ile Pro Glu
130 135 140

Gly Met Gln Cys Ser Cys Gly Ile Asp Tyr Tyr Thr Pro His Glu Glu
145 150 155 160

Thr Asn Asn Glu Ser Phe Val Ile Tyr Met Phe Val Val His Phe Ile
165 170 175

Ile Pro Leu Ile Val Ile Phe Phe Cys Tyr Gly Gln Leu Val Phe Thr
180 185 190

Val Lys Glu Ala Ala Ala Ser Ala Thr Thr Gln Lys Ala Glu Lys Glu
195 200 205

Val Thr Arg Met Val Ile Ile Met Val Ile Ala Phe Leu Ile Cys Trp
210 215 220

Leu Pro Tyr Ala Gly Val Ala Phe Tyr Ile Phe Thr His Gln Gly Ser
225 230 235 240

Asp Phe Gly Pro Ile Phe Met Thr Ile Pro Ala Phe Phe Ala Lys Thr
245 250 255

Ser Ala Val Tyr Asn Pro Val Ile Tyr Ile Met
260 265

<210> 18

<211> 214

<212> PRT

<213> Halobacterium salinarum

<400> 18

Gly Arg Pro Glu Trp Ile Trp Leu Ala Leu Gly Thr Ala Leu Met Gly
1 5 10 15

Leu Gly Thr Leu Tyr Phe Leu Val Lys Gly Met Gly Val Ser Asp Pro
20 25 30

Asp Ala Lys Lys Phe Tyr Ala Ile Thr Thr Leu Val Pro Ala Ile Ala
35 40 45

Phe Thr Met Tyr Leu Ser Met Leu Leu Gly Tyr Gly Leu Thr Met Val
50 55 60

Pro Phe Gly Gly Glu Gln Asn Pro Ile Tyr Trp Ala Arg Tyr Ala Asp
65 70 75 80

Trp Leu Phe Thr Thr Pro Leu Leu Leu Leu Asp Leu Ala Leu Leu Val
85 90 95

Asp Ala Asp Gln Gly Thr Ile Leu Ala Leu Val Gly Ala Asp Gly Ile
100 105 110

Met Ile Gly Thr Gly Leu Val Gly Ala Leu Thr Lys Val Tyr Ser Tyr
115 120 125

Arg Phe Val Trp Trp Ala Ile Ser Thr Ala Ala Met Leu Tyr Ile Leu
130 135 140

Tyr Val Leu Phe Phe Gly Phe Ser Met Arg Pro Glu Val Ala Ser Thr
145 150 155 160

Phe Lys Val Leu Arg Asn Val Thr Val Val Leu Trp Ser Ala Tyr Pro
165 170 175

Val Val Trp Leu Ile Gly Ser Glu Gly Ala Gly Ile Val Pro Leu Asn
180 185 190

Ile Glu Thr Leu Leu Phe Met Val Leu Asp Val Ser Ala Lys Val Gly
195 200 205

Phe Gly Leu Ile Leu Leu
210

<210> 19
<211> 246
<212> PRT
<213> Rhodopseudomonas viridis

<400> 19

Gly Thr Leu Ile Gly Gly Asp Leu Phe Asp Phe Trp Val Gly Pro Tyr
1 5 10 15

Phe Val Gly Phe Phe Gly Val Ser Ala Ile Phe Phe Ile Phe Leu
20 25 30

Gly Val Ser Leu Ile Gly Tyr Ala Ala Ser Gln Gly Pro Thr Trp Asp
35 40 45

Pro Phe Ala Ile Ser Ile Asn Pro Pro Asp Leu Lys Tyr Gly Leu Gly
50 55 60

Ala Ala Pro Leu Leu Glu Gly Gly Phe Trp Gln Ala Ile Thr Val Cys
65 70 75 80

Ala Leu Gly Ala Phe Ile Ser Trp Met Leu Arg Glu Val Glu Ile Ser
85 90 95

Arg Lys Leu Gly Ile Gly Trp His Val Pro Leu Ala Phe Cys Val Pro
100 105 110

Ile Phe Met Phe Cys Val Leu Gln Val Phe Arg Pro Leu Leu Gly
115 120 125

Ser Trp Gly His Ala Phe Pro Tyr Gly Ile Leu Ser His Leu Asp Trp
130 135 140

Val Asn Asn Phe Gly Tyr Gln Tyr Leu Asn Trp His Tyr Asn Pro Gly
145 150 155 160

His Met Ser Ser Val Ser Phe Leu Phe Val Asn Ala Met Ala Leu Gly
165 170 175

Leu His Gly Gly Leu Ile Leu Ser Val Ala Asn Pro Gly Asp Gly Asp
180 185 190

Lys Val Lys Thr Ala Glu His Glu Asn Gln Tyr Phe Arg Asp Val Val
195 200 205

Gly Tyr Ser Ile Gly Ala Leu Ser Ile His Arg Leu Gly Leu Phe Leu
210 215 220

Ala Ser Asn Ile Phe Leu Thr Gly Ala Phe Gly Thr Ile Ala Ser Gly
225 230 235 240

Pro Phe Trp Thr Arg Gly
245

<210> 20
<211> 259
<212> PRT
<213> Rhodopseudomonas viridis

<400> 20

Tyr Ser Tyr Trp Leu Gly Lys Ile Gly Asp Ala Gln Ile Gly Pro Ile
1 5 10 15

Tyr Leu Gly Ala Ser Gly Ile Ala Ala Phe Ala Phe Gly Ser Thr Ala
20 25 30

Ile Leu Ile Ile Leu Phe Asn Met Ala Ala Glu Val His Phe Asp Pro
35 40 45

Leu Gln Phe Phe Arg Gln Phe Phe Trp Leu Gly Leu Tyr Pro Pro Lys
50 55 60

Ala Gln Tyr Gly Met Gly Ile Pro Pro Leu His Asp Gly Gly Trp Trp
65 70 75 80

Leu Met Ala Gly Leu Phe Met Thr Leu Ser Leu Gly Ser Trp Trp Ile
85 90 95

Arg Val Tyr Ser Arg Ala Arg Ala Leu Gly Leu Gly Thr His Ile Ala
100 105 110

Trp Asn Phe Ala Ala Ala Ile Phe Phe Val Leu Cys Ile Gly Cys Ile
115 120 125

His Pro Thr Leu Val Gly Ser Trp Ser Glu Gly Val Pro Phe Gly Ile
130 135 140

Trp Pro His Ile Asp Trp Leu Thr Ala Phe Ser Ile Arg Tyr Gly Asn
145 150 155 160

Phe Tyr Tyr Cys Pro Trp His Gly Phe Ser Ile Gly Phe Ala Tyr Gly
165 170 175

Cys Gly Leu Leu Phe Ala Ala His Gly Ala Thr Ile Leu Ala Val Ala
180 185 190

Arg Phe Gly Gly Asp Arg Glu Ile Glu Gln Ile Thr Asp Arg Gly Thr
195 200 205

Ala Val Glu Arg Ala Ala Leu Phe Trp Arg Trp Thr Ile Gly Phe Asn
210 215 220

Ala Thr Ile Glu Ser Val His Arg Trp Gly Trp Phe Phe Ser Leu Met
225 230 235 240

Val Met Val Ser Ala Ser Val Gly Ile Leu Leu Thr Gly Thr Phe Val
245 250 255

Asp Asn Trp

<210> 21
<211> 145
<212> PRT
<213> Escherichia coli

<400> 21

Thr Val Thr Gly Gly Tyr Ala Gln Ser Asp Ala Gln Gly Gln Met Asn
1 5 10 15

Lys Met Gly Gly Phe Asn Leu Lys Tyr Arg Tyr Glu Glu Asp Asn Ser
20 25 30

Pro Leu Gly Val Ile Gly Ser Phe Thr Tyr Thr Glu Lys Ser Arg Thr
35 40 45

Ala Ser Ser Gly Asp Tyr Asn Lys Asn Gln Tyr Tyr Gly Ile Thr Ala
50 55 60

Gly Pro Ala Tyr Arg Ile Asn Asp Trp Ala Ser, Ile Tyr Gly Val Val
65 70 75 80

Gly Val Gly Tyr Gly Lys Phe Gln Thr Thr Glu Tyr Pro Thr Tyr Lys
85 90 95

Asn Asp Thr Ser Asp Tyr Gly Phe Ser Tyr Gly Ala Gly Leu Gln Phe
100 105 110

Asn Pro Met Glu Asn Val Ala Leu Asp Phe Ser Tyr Glu Gln Ser Arg
115 120 125

Ile Arg Ser Val Asp Val Gly Thr Trp Ile Ala Gly Val Gly Tyr Arg
130 135 140

Phe
145

<210> 22

<211> 153
<212> PRT
<213> Escherichia coli

<400> 22

Tyr His Asp Thr Gly Leu Ile Asn Asn Asn Gly Pro Thr His Glu Asn
1 5 10 15

Lys Leu Gly Ala Gly Ala Phe Gly Gly Tyr Gln Val Asn Pro Tyr Val
20 25 30

Gly Phe Glu Met Gly Tyr Asp Trp Leu Gly Arg Met Pro Tyr Lys Gly
35 40 45

Ser Val Glu Asn Gly Ala Tyr Lys Ala Gln Gly Val Gln Leu Thr Ala
50 55 60

Lys Leu Gly Tyr Pro Ile Thr Asp Asp Leu Asp Ile Tyr Thr Arg Leu
65 70 75 80

Gly Gly Met Val Trp Arg Ala Asp Thr Tyr Ser Asn Val Tyr Gly Lys
85 90 95

Asn His Asp Thr Gly Val Ser Pro Val Phe Ala Gly Gly Val Glu Tyr
100 105 110

Ala Ile Thr Pro Glu Ile Ala Thr Arg Leu Glu Tyr Gln Trp Thr Asn
115 120 125

Asn Ile Gly Asp Ala His Thr Ile Gly Thr Arg Pro Asp Asn Gly Met
130 135 140

Leu Ser Leu Gly Val Ser Tyr Arg Phe
145 150

<210> 23
<211> 301
<212> PRT
<213> Unknown

<220>

<223> Porin Crystal Structure B, 102mA, Weiss, M.S., Schultz, G.E., Structure of Porin Refined at 1.8A Resolution, 277 J. Mol. Bio. 493 (1992)

<400> 23

Glu Val Lys Leu Ser Gly Asp Ala Arg Met Gly Val Met Tyr Asn Gly
1 5 10 15

Asp Asp Trp Asn Phe Ser Ser Arg Ser Arg Val Leu Phe Thr Met Ser
20 25 30

Gly Thr Thr Asp Ser Gly Leu Glu Phe Gly Ala Ser Phe Lys Ala His
35 40 45

Glu Ser Val Gly Ala Glu Thr Gly Glu Asp Gly Thr Val Phe Leu Ser
50 55 60

Gly Ala Phe Gly Lys Ile Glu Met Gly Asp Ala Leu Gly Ala Ser Glu
65 70 75 80

Ala Leu Phe Gly Asp Leu Tyr Glu Val Gly Tyr Thr Asp Leu Asp Asp
85 90 95

Arg Gly Gly Asn Asp Ile Pro Tyr Leu Thr Gly Asp Glu Arg Leu Thr
100 105 110

Ala Glu Asp Asn Pro Val Leu Leu Tyr Thr Tyr Ser Ala Gly Ala Phe
115 120 125

Ser Val Ala Ala Ser Met Ser Asp Gly Lys Val Gly Glu Thr Ser Glu
130 135 140

Asp Asp Ala Gln Glu Met Ala Val Ala Ala Tyr Thr Phe Gly Asn
145 150 155 160

Tyr Thr Val Gly Leu Gly Tyr Glu Lys Ile Asp Ser Pro Asp Thr Ala
165 170 175

Leu Met Ala Asp Met Glu Gln Leu Glu Leu Ala Ala Ile Ala Lys Phe
180 185 190

Gly Ala Thr Asn Val Lys Ala Tyr Tyr Ala Asp Gly Glu Leu Asp Arg
195 200 205

Asp Phe Ala Arg Ala Val Phe Asp Leu Thr Pro Val Ala Ala Ala Ala
210 215 220

Thr Ala Val Asp His Lys Ala Tyr Gly Leu Ser Val Asp Ser Thr Phe
225 230 235 240

Gly Ala Thr Thr Val Gly Gly Tyr Val Gln Val Leu Asp Ile Asp Thr
245 250 255

Ile Asp Asp Val Thr Tyr Gly Leu Gly Ala Ser Tyr Asp Leu Gly
260 265 270

Gly Gly Ala Ser Ile Val Gly Gly Ile Ala Asp Asn Asp Leu Pro Asn
275 280 285

Ser Asp Met Val Ala Asp Leu Gly Val Lys Phe Lys Phe
290 295 300

<210> 24
<211> 322
<212> PRT
<213> Klebsiella pneumoniae

<400> 24

Lys Leu Asp Leu Tyr Gly Lys Ile Asp Gly Leu His Tyr Phe Ser Asp
1 5 10 15

Asp Lys Asp Val Asp Gly Asp Gln Thr Tyr Met Arg Leu Gly Val Lys
20 25 30

Gly Glu Thr Gln Ile Asn Asp Gln Leu Thr Gly Tyr Gly Gln Trp Glu
35 40 45

Tyr Asn Val Gln Ala Asn Asn Thr Glu Ser Ser Asp Gln Ala Trp
50 55 60

Thr Arg Leu Ala Phe Ala Gly Leu Lys Phe Gly Asp Ala Gly Ser Phe
65 70 75 80

Asp Tyr Gly Arg Asn Tyr Gly Val Val Tyr Asp Val Thr Ser Trp Thr
85 90 95

Asp Val Leu Pro Glu Phe Gly Asp Thr Tyr Gly Ser Asp Asn Phe
100 105 110

Leu Gln Ser Arg Ala Asn Gly Val Ala Thr Tyr Arg Asn Ser Asp Phe
115 120 125

Phe Gly Leu Val Gly Leu Asn Phe Ala Leu Gln Tyr Gln Gly Lys Asn
130 135 140

Gly Ser Val Ser Gly Glu Gly Ala Thr Asn Asn Gly Arg Gly Ala Leu
145 150 155 160

Lys Gln Asn Gly Asp Gly Phe Gly Thr Ser Val Thr Tyr Asp Ile Phe
165 170 175

Asp Gly Ile Ser Ala Gly Phe Ala Tyr Ala Asn Ser Lys Arg Thr Asp
180 185 190

Asp Gln Asn Gln Leu Leu Leu Gly Glu Gly Asp His Ala Glu Thr Tyr
195 200 205

Thr Gly Gly Leu Lys Tyr Asp Ala Asn Asn Ile Tyr Leu Ala Thr Gln
210 215 220

Tyr Thr Gln Thr Tyr Asn Ala Thr Arg Ala Gly Ser Leu Gly Phe Ala
225 230 235 240

Asn Lys Ala Gln Asn Phe Glu Val Ala Ala Gln Tyr Gln Phe Asp Phe
245 250 255

Gly Leu Arg Pro Ser Val Ala Tyr Leu Gln Ser Lys Gly Lys Asp Leu
260 265 270

Asn Gly Tyr Gly Asp Gln Asp Ile Leu Lys Tyr Val Asp Val Gly Ala
275 280 285

Thr Tyr Tyr Phe Asn Lys Asn Met Ser Thr Tyr Val Asp Tyr Lys Ile
290 295 300

Asn Leu Leu Asp Asp Asn Ser Phe Thr Arg Ser Ala Gly Ile Ser Thr
305 310 315 320

Asp Asp

<210> 25
<211> 420
<212> PRT

<213> *Salmonella typhimurium*

<400> 25

Asp Phe His Gly Tyr Ala Arg Ser Gly Ile Gly Trp Thr Gly Ser Gly
1 5 10 15

Gly Glu Gln Gln Cys Phe Gln Ala Thr Gly Ala Gln Ser Lys Tyr Arg
20 25 30

Leu Gly Asn Glu Cys Glu Thr Tyr Ala Glu Leu Lys Leu Gly Gln Glu
35 40 45

Val Trp Lys Glu Gly Asp Lys Ser Phe Tyr Phe Asp Thr Asn Val Ala
50 55 60

Tyr Ser Val Asn Gln Gln Asn Asp Trp Glu Ser Thr Asp Pro Ala Phe
65 70 75 80

Arg Glu Ala Asn Val Gln Gly Lys Asn Leu Ile Glu Trp Leu Pro Gly
85 90 95

Ser Thr Ile Trp Ala Gly Lys Arg Phe Tyr Gln Arg His Asp Val His
100 105 110

Met Ile Asp Phe Tyr Tyr Trp Asp Ile Ser Gly Pro Gly Ala Gly Ile
115 120 125

Glu Asn Ile Asp Leu Gly Phe Gly Lys Leu Ser Leu Ala Ala Thr Arg
130 135 140

Ser Thr Glu Ala Gly Gly Ser Tyr Thr Phe Ser Ser Gln Asn Ile Tyr
145 150 155 160

Asp Glu Val Lys Asp Thr Ala Asn Asp Val Phe Asp Val Arg Leu Ala
165 170 175

Gly Leu Gln Thr Asn Pro Asp Gly Val Leu Glu Leu Gly Val Asp Tyr
180 185 190

Gly Arg Ala Asn Thr Thr Asp Gly Tyr Lys Leu Ala Asp Gly Ala Ser
195 200 205

Lys Asp Gly Trp Met Phe Thr Ala Glu His Thr Gln Ser Met Leu Lys

210

215

220

Gly Tyr Asn Lys Phe Val Val Gln Tyr Ala Thr Asp Ala Met Thr Thr
225 230 235 240

Gln Gly Lys Gly Gln Ala Arg Gly Ser Asp Gly Ser Ser Ser Phe Thr
245 250 255

Glu Lys Ile Asn Tyr Ala Asn Lys Val Ile Asn Asn Asn Gly Asn Met
260 265 270

Trp Arg Ile Leu Asp His Gly Ala Ile Ser Leu Gly Asp Lys Trp Asp
275 280 285

Leu Met Tyr Val Gly Met Tyr Gln Asn Ile Asp Trp Asp Asn Asn Leu
290 295 300

Gly Thr Glu Trp Trp Thr Val Gly Val Arg Pro Met Tyr Lys Trp Thr
305 310 315 320

Pro Ile Met Ser Thr Leu Leu Glu Val Gly Tyr Asp Asn Val Lys Ser
325 330 335

Gln Gln Thr Gly Asp Arg Asn Asn Gln Tyr Lys Ile Thr Leu Ala Gln
340 345 350

Gln Trp Gln Ala Gly Asp Ser Ile Trp Ser Arg Pro Ala Ile Arg Ile
355 360 365

Phe Ala Thr Tyr Ala Lys Trp Asp Glu Lys Trp Gly Tyr Ile Lys Asp
370 375 380

Gly Asp Asn Ile Ser Arg Tyr Ala Ala Ala Thr Asn Ser Gly Ile Ser
385 390 395 400

Thr Asn Ser Arg Gly Asp Ser Asp Glu Trp Thr Phe Gly Ala Gln Met
405 410 415

Glu Ile Trp Trp
420

<210> 26
<211> 410

<212> PRT

<213> Salmonella typhimurium

<400> 26

Glu Phe His Gly Tyr Ala Arg Ser Gly Val Ile Met Asn Asp Ser Gly
1 5 10 15

Ala Ser Thr Lys Ser Gly Ala Tyr Ile Thr Pro Ala Gly Glu Thr Gly
20 25 30

Gly Ala Ile Gly Arg Leu Gly Asn Gln Ala Asp Thr Tyr Val Glu Met
35 40 45

Asn Leu Glu His Lys Gln Thr Leu Asp Asn Gly Ala Thr Thr Arg Phe
50 55 60

Lys Val Met Val Ala Asp Gly Gln Thr Ser Tyr Asn Asp Trp Thr Ala
65 70 75 80

Ser Thr Ser Asp Leu Asn Val Arg Gln Ala Phe Val Glu Leu Gly Asn
85 90 95

Leu Pro Thr Phe Ala Gly Pro Phe Lys Gly Ser Thr Leu Trp Ala Gly
100 105 110

Lys Arg Phe Asp Arg Asp Asn Phe Asp Ile His Trp Ile Asp Ser Asp
115 120 125

Val Val Phe Leu Ala Gly Thr Gly Gly Ile Tyr Asp Val Lys Trp
130 135 140

Asn Asp Gly Leu Arg Ser Asn Phe Ser Leu Tyr Gly Arg Asn Phe Gly
145 150 155 160

Asp Ile Asp Asp Ser Ser Asn Ser Val Gln Asn Tyr Ile Leu Thr Met
165 170 175

Asn His Phe Ala Gly Pro Leu Gln Met Met Val Ser Gly Leu Arg Ala
180 185 190

Lys Asp Asn Asp Glu Arg Lys Asp Ser Asn Gly Asn Leu Ala Lys Gly
195 200 205

Asp Ala Ala Asn Thr Gly Val His Ala Leu Leu Gly Leu His Asn Asp
210 215 220

Ser Phe Tyr Gly Leu Arg Asp Gly Ser Ser Lys Thr Ala Leu Leu Tyr
225 230 235 240

Gly His Gly Leu Gly Ala Glu Val Lys Gly Ile Gly Ser Asp Gly Ala
245 250 255

Leu Arg Pro Gly Ala Asp Thr Trp Arg Ile Ala Ser Tyr Gly Thr Thr
260 265 270

Pro Leu Ser Glu Asn Trp Ser Val Ala Pro Ala Met Leu Ala Gln Arg
275 280 285

Ser Lys Asp Arg Tyr Ala Asp Gly Asp Ser Tyr Gln Trp Ala Thr Phe
290 295 300

Asn Leu Arg Leu Ile Gln Ala Ile Asn Gln Asn Phe Ala Leu Ala Tyr
305 310 315 320

Glu Gly Ser Tyr Gln Tyr Met Asp Leu Lys Pro Glu Gly Tyr Asn Asp
325 330 335

Arg Gln Ala Val Asn Gly Ser Phe Tyr Lys Leu Thr Phe Ala Pro Thr
340 345 350

Phe Lys Val Gly Ser Ile Gly Asp Phe Phe Ser Arg Pro Glu Ile Arg
355 360 365

Phe Tyr Thr Ser Trp Met Asp Trp Ser Lys Lys Leu Asn Asn Tyr Ala
370 375 380

Ser Asp Asp Ala Leu Gly Ser Asp Gly Phe Asn Ser Gly Gly Glu Trp
385 390 395 400

Ser Phe Gly Val Gln Met Glu Thr Trp Phe
405 410